

A Distance Measure Comparison to Improve Crowding in Multi- Modal Optimization Problems.

**D. Todd Vollmer, Idaho National Laboratory
Terrence Soule, University of Idaho
Milos Manic, University of Idaho**

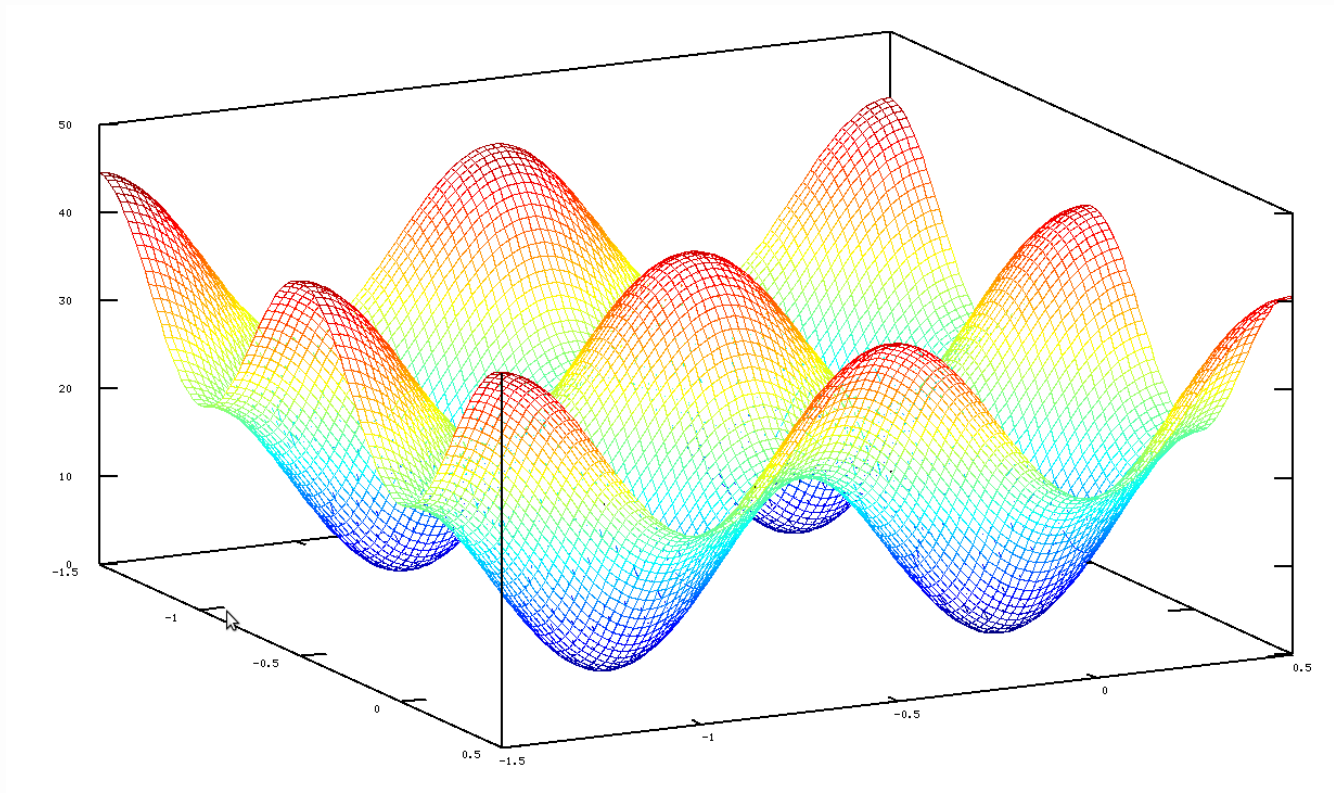
August 10-12, 2010
Idaho Falls, Idaho

What is this?



Problem Space

- There are problems in which a number of points are potentially good solutions, local optima, while not necessarily being a global best answer.
- Multi-modal optimization problems are of interest to researchers solving real world problems in areas such as control systems and power engineering tasks.



Genetic Algorithms

- A Genetic Algorithm (GA) is a heuristic search technique inspired by concepts of evolutionary biology.
 - Population – An initially randomly generated set of possible solutions.
 - Variation Operators – mutation and crossover.
 - Fitness Function – evaluation of a solution.
 - Selection and Replacement operators.
- Conventional GA's tend to converge to just one optima.
- A 2008 review of papers mainly in IEEE Transactions and IEE proceedings found ~1000 papers dealing with power engineering and GA's. (N. Rajkumar, et al.)

GA Variations

- Deterministic Crowding (DC)
 - After crossover and mutation, each resulting new solution replaces the most *similar* parent used to create it if the new solution has a higher fitness value.
- Restricted Tournament Selection (RTS)
 - The new candidate solutions compete with a fixed number of randomly chosen individuals (called a Crowding Factor) from the population. The individual from the CF that is *closest* to a given new solution competes with that solution for survival.

How do we determine *similarity* and *closeness*? A Euclidean distance measure is common and used in a great variety of algorithms. How does it compare with a Mahalanobis distance measure when utilizing DC and RTS.

Euclidean Distance

- Simple and familiar.
- Potential issues with scale and correlation.

$$d(x, y) = \sum_{i=1}^n \sqrt{(x_i - y_i)^2}$$

Mahalanobis

It is based on correlations between variables by which different patterns can be identified and analyzed. It is a useful way of determining similarity of an unknown sample set or point to a known one. It differs from Euclidean distance in that it takes into account the correlations of the data set and is scale-invariant, i.e. not dependent on the scale of measurements.

$$x = (x_1, x_2, x_3, \dots, x_N)^T$$

$$\mu = (\mu_1, \mu_2, \mu_3, \dots, \mu_N)^T$$

$$D_M(x) = \sqrt{(x - \mu)^T S^{-1} (x - \mu)}.$$

What is S^{-1} ?

Covariance Matrix

What is S^{-1} ? This is the inverse covariance matrix. The covariance is always calculated between 2 dimensions. Covariance is a measure of how much the dimensions vary from the mean with respect to each other. If we have a dataset with more than 2 dimensions there are several covariance calculations that can be performed.

Ex. 3 dimensions (x,y,z)

cov (x,y)

cov (x,z)

cov (y,z)

Given n dimensions we can calculate them all and put them in a matrix.

$$C^{n \times n} = (c_{i,j}, c_{i,j} = \text{cov}(\text{Dim}_i, \text{Dim}_j))$$

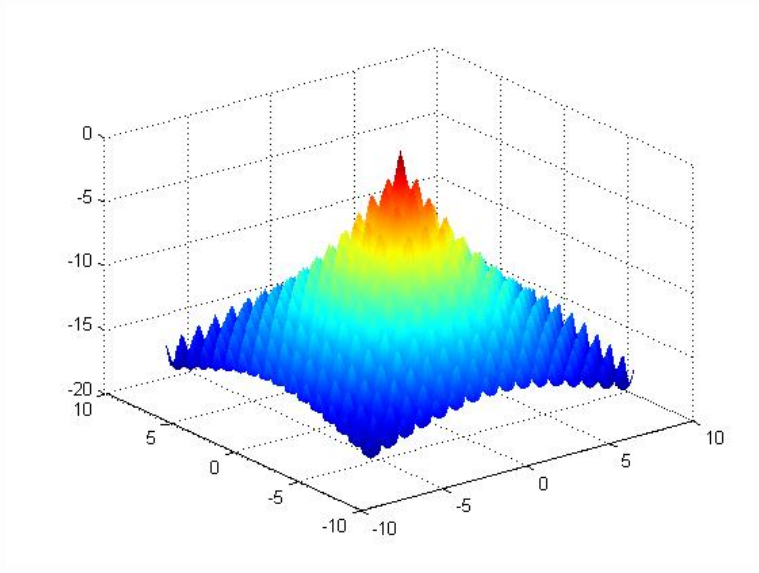
Covariance Matrix Reloaded

$$C = \begin{pmatrix} \text{cov}(x, x) & \text{cov}(x, y) & \text{cov}(x, z) \\ \text{cov}(y, x) & \text{cov}(y, y) & \text{cov}(y, z) \\ \text{cov}(z, x) & \text{cov}(z, y) & \text{cov}(z, z) \end{pmatrix}$$

Example in 3 dimensions. Note that the covariance of a variable with itself is just the variance (diagonal).

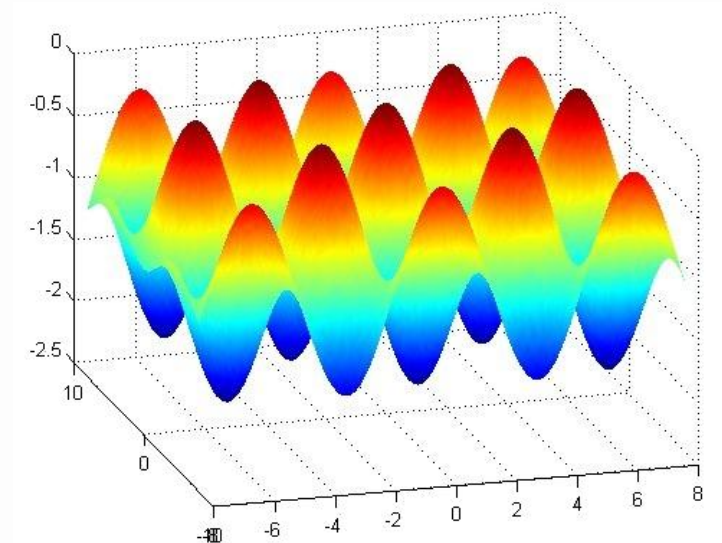
Also since $\text{cov}(a,b) == \text{cov}(b,a)$ this matrix is symmetrical about the main diagonal. Finally this is a square matrix (nxn).

Problem Set



$$F_{ak}(x) = 20 + e - 20 \exp \left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \right) - \exp \left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i) \right)$$

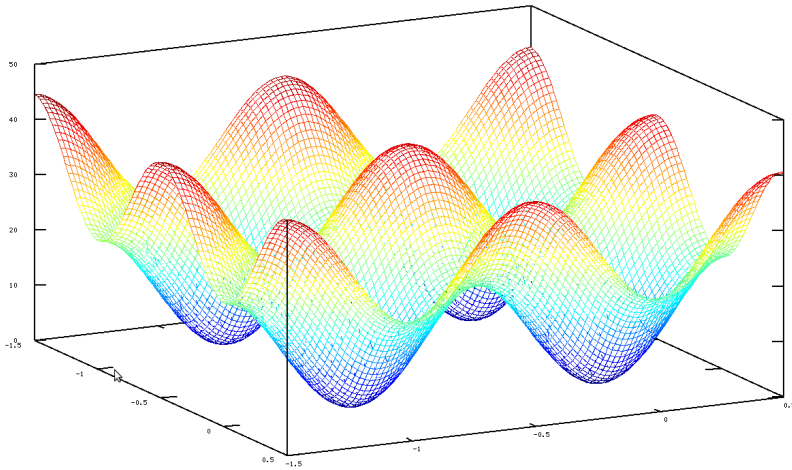
Ackley



$$F_g(x) = 1 + \sum_{i=1}^n \frac{x_i^2}{400} - \prod_{i=1}^n \cos \left(\frac{x_i}{\sqrt{i}} \right)$$

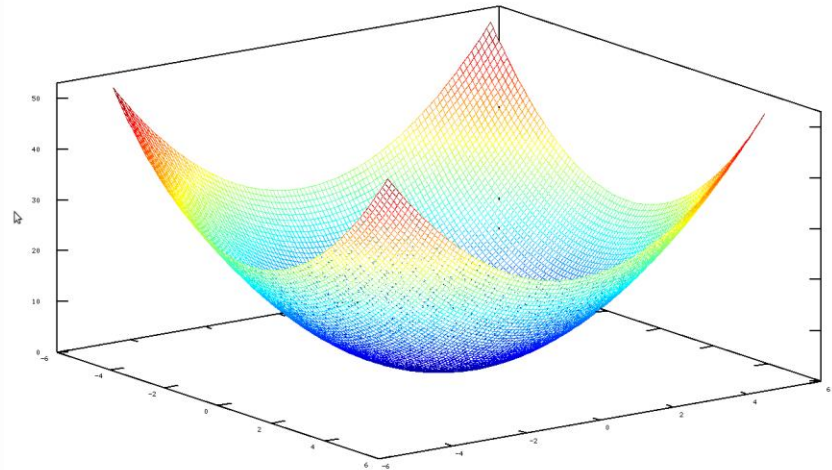
Griewank

Problem Set



$$F_{ra}(x) = nA + \sum x_i^2 - A \cos(wx_i)$$

Rastrigin



$$F_s(x) = \sum_{i=1}^n x_i^2$$

Sphere

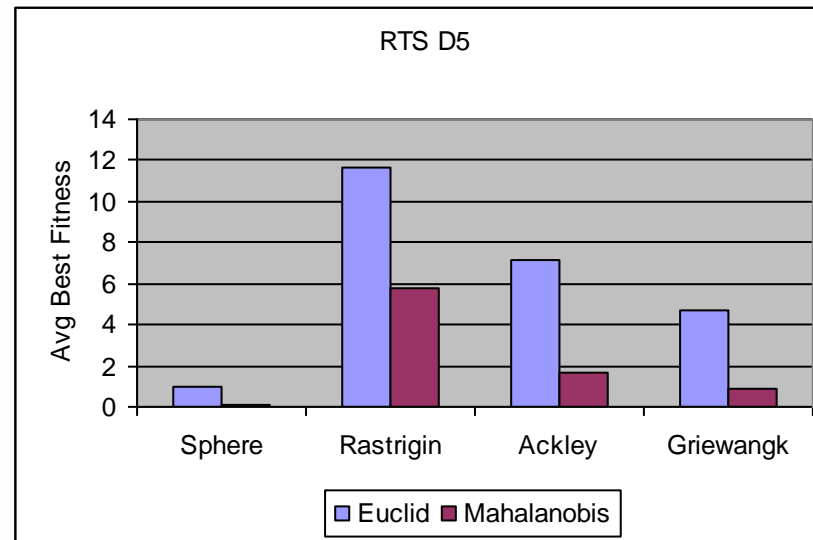
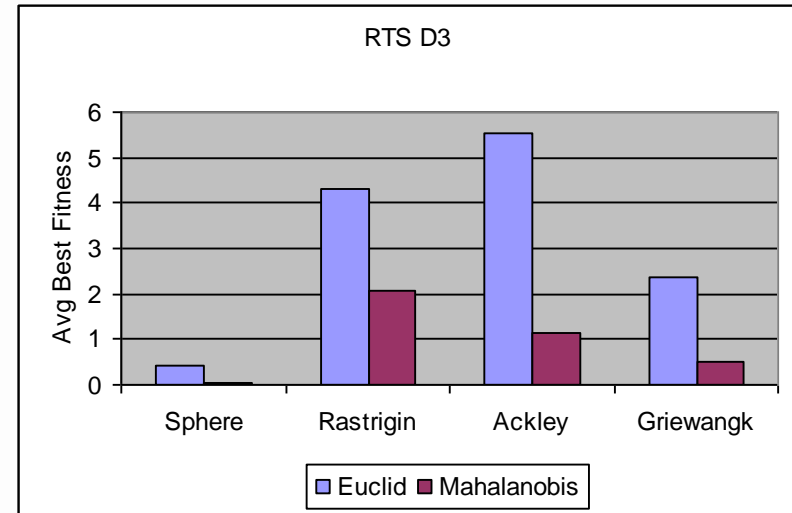
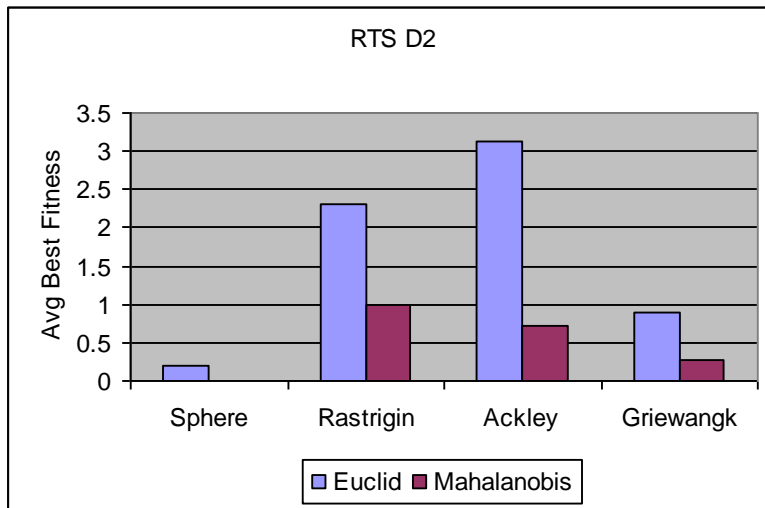
GA Parameters

	Sphere	Rastrigin	Ackley	Griewangk	M6
Iteration ²	400	400	400	400	400
Iteration ³	500	500	500	500	-
Iteration ⁵	600	600	600	600	-
Optima ²	1	4	1	5	25
Optima ³	1	8	1	5	-
Optima ⁵	1	32	1	5	-
niche	0.2	0.1	1	0.9	0.5

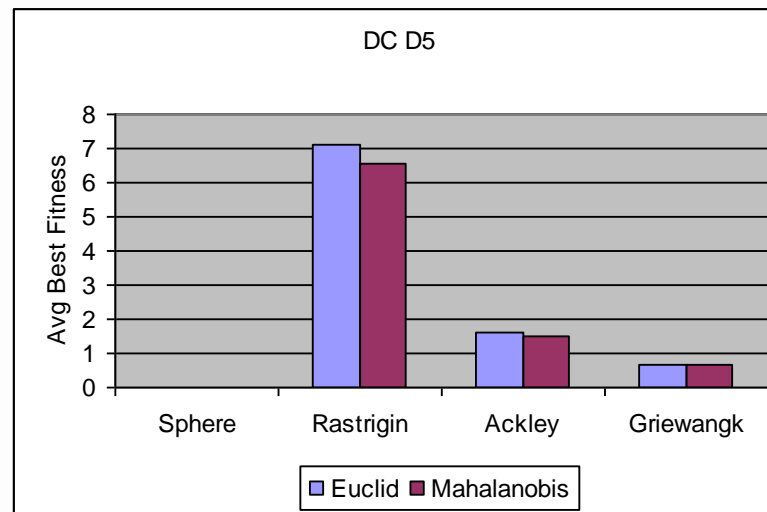
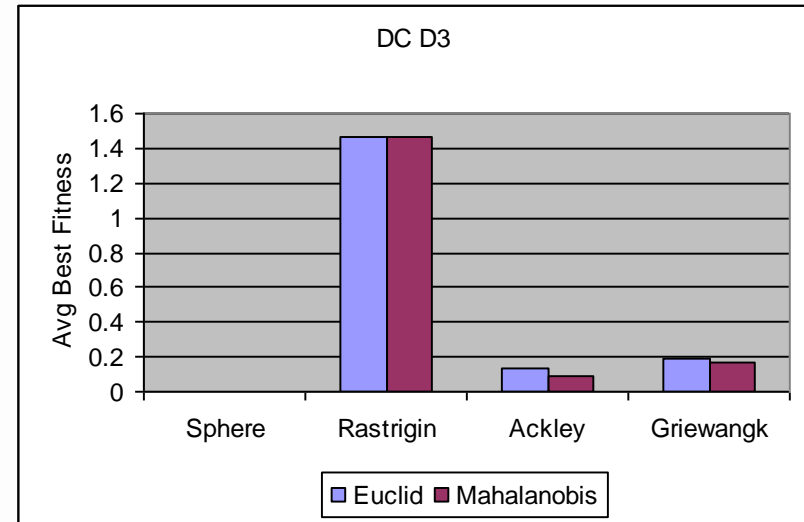
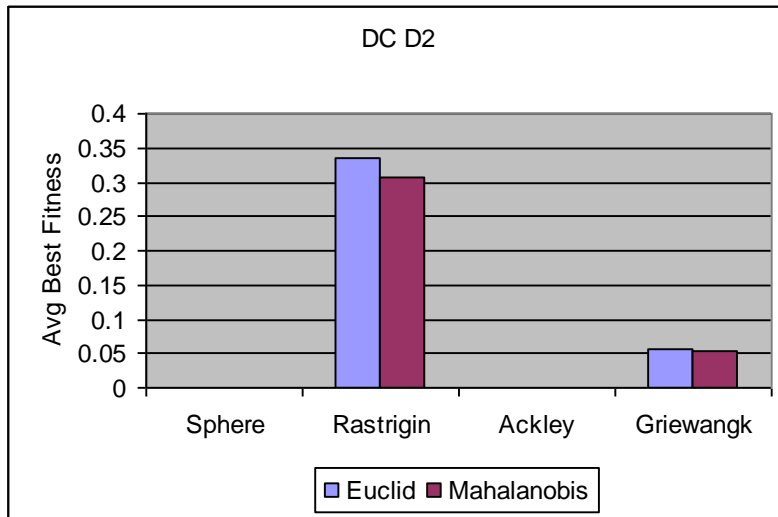
Global Optimum Results

	DC		RTS		
<i>Functions</i>	<i>Euclid</i>	<i>Mahal</i>	<i>Euclid</i>	<i>Mahal</i>	<i>Dim.</i>
Sphere	100	100	97	100	2
Rastrigin	98	98	91	89	2
Ackley	100	100	70	95	2
Griewangk	20	11	0	1	2
M6	3	2	3	4	2
Sphere	100	100	61	100	3
Rastrigin	49	42	35	35	3
Ackley	98	93	17	73	3
Griewangk	0	1	0	0	3
Sphere	100	100	5	97	5
Rastrigin	0	1	1	3	5
Ackley	35	36	0	33	5
Griewangk	0	0	0	0	5

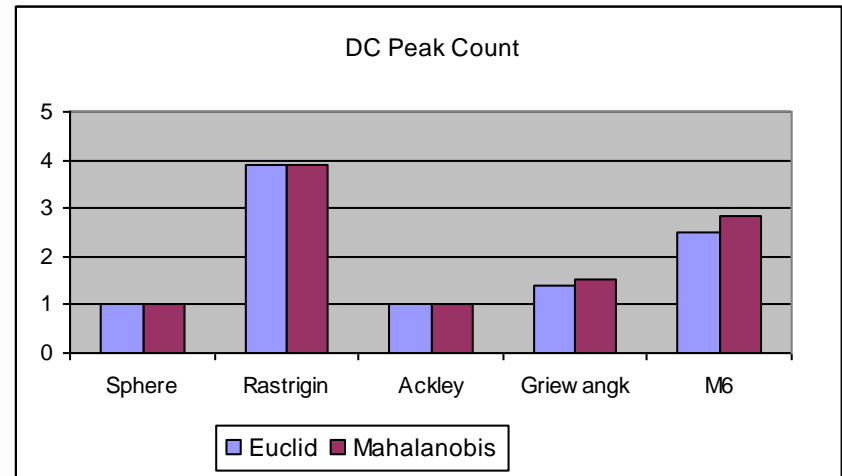
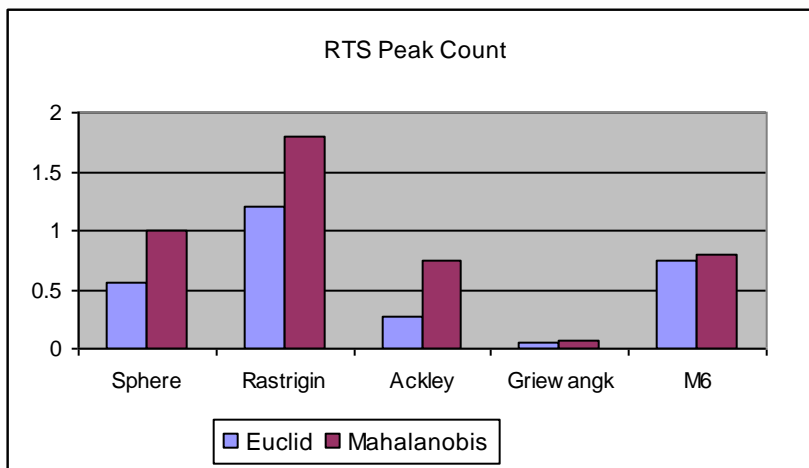
RTS Average Best Fitness



DC Average Best Fitness



Peak Counts



Conclusion

- For DC there is little to no difference between distance measures with the possible exception of global optima.
- RTS consistently showed improvement using Mahalanobis in all three quality measures.

Good Genetics – Idaho 2009 Record



Great Genetics – World Record 2009

